Scaling patterns inform ontogenetic transitions away from cleaning in *Thalassoma* wrasses

Vikram B. Baliga[1*], Rita S. Mehta[1]

[1]*Department of Ecology and Evolutionary Biology, Long Marine Laboratory, University of California Santa Cruz, Santa Cruz, CA 95060, USA*

***Correspondence:**

Vikram B. Baliga

Department of Ecology and Evolutionary Biology

University of California Santa Cruz

Santa Cruz, CA 95060

Email: vbaliga@ucsc.edu

Summary: In fishes, cleaning is a mutualistic behavior wherein a species will remove and consume ectoparasites from other organisms. More than two thirds of cleaner fishes display this behavior predominately as juveniles and discontinue cleaning as adults; such species are here referred to as “facultative cleaners”. Whether allometric changes in morphological traits coincide with ontogenetic shifts away from cleaning is unknown. We tested the hypothesis that transitions away from cleaning are associated with scaling patterns in the feeding apparatus of facultative cleaners, and then compared such patterns to those exhibited by non-cleaner congeners. We measured functional traits related to feeding, such as vertical gape distance, premaxillary protrusion distance, and maxillary kinematic transmission coefficient (KT) in each ontogenetic series of 11 *Thalassoma* wrasses (Labridae). As these fishes predominately capture prey via biting, we modeled bite force using MandibLever (v3.3) to create an ontogenetic trajectory of bite force for each species. Our results indicate that cleaner fishes in *Thalassoma*, as juveniles, possess jaws with low mobility that exhibit weaker bite forces. Additionally, there was remarkable consistency in the range of body lengths over which we observed significant differences between facultative cleaners and non-cleaners in bite force, vertical gape distance, and premaxillary protrusion distance. Through ontogeny, facultative cleaner fishes exhibit positive allometry for a number of functionally important feeding traits, which possibly obviates the need to continue cleaning.

Key Words: cleaning, scaling, functional morphology, feeding, bite force, maxillary KT, wrasses

Introduction

The consequences of size on the structure and function of organismal systems are pervasive (McMahon, 1984; Schmidt-Nielson, 1984). Therefore, it is not surprising that a central focus of functional morphology studies is to understand how the scaling of the musculoskeletal system influences the scaling of functional traits across ontogeny. Previous studies have argued that ontogenetic shifts in ecology often drive adaptive changes in the scaling of musculoskeletal systems resulting in differential performance (McMahon, 1984; Richard and Wainwright, 1995; Deban and O’Reilly, 2005; Herrell and Gibb, 2006; Pfaller et al., 2011). These studies, in turn, support the idea that allometric changes in morphology often co-occur with changes in feeding strategies, locomotor behavior, or habitat use. Fewer studies, however, have compared the ontogenetic trajectories of functional traits across closely-related species to better understand the extent to which such patterns of scaling may be specifically adaptive during a particular life history stage (but see Mitteroecker et al., 2004; Herrel and O’Reilly, 2006; Frédérich and Sheets, 2009; Wilson and Sánchez-Villagra, 2010).

Cleaning behavior provides us with the opportunity to examine both the patterns of scaling of the musculoskeletal system and shifts in feeding ecology within a clade of tropical reef fishes. In fishes, cleaning is a mutualistic behavior wherein an individual consumes ectoparasites (generally juvenile gnathiid or cymothoid isopods) off other organisms. Over 120 species of teleost fishes, from diverse lineages including wrasses, cichlids, surfperches, and gobies, exhibit this behavior (Coté, 2000; Froese and Pauly, 2014). The evolution of cleaning in fishes is especially fascinating as it represents one of the few examples of mutualistic behavior between vertebrate species (Bronstein, 1994; Poulin and Grutter, 1996). This behavior relies not only on the ability of client species to recognize cleaners but also requires that cleaners possess morphological, functional, and behavioral traits that are necessary to find and remove ectoparasites.

Over two thirds of fishes that clean do so predominately as juveniles (Froese and Pauly, 2014), exhibiting ontogenetic transitions away from cleaning behavior. While these species are referred to as “facultative (juvenile) cleaners” in the literature (Coté, 2000), for simplicity we here refer to these species as “facultative” cleaners.

In some cases, cleaners may also consume the mucus coating of their clients. This behavior has been observed in several species including *Labroides* cleaner wrasses, as well as *Elacatinus* cleaner gobies (Gorlick, 1980; Grutter, 1997; Bshary and Grutter, 2002; Grutter and Bshary, 2004; Soares et al., 2009; Soares et al., 2010). Notably, in both of these clades, cleaners are described as “obligate”, indicating that they obtain, on average, 85% of dietary items through cleaning (Coté, 2000). Here, mucivory adds an important dimension to the dynamics exhibited between cleaners and their clients. Clients appear to prefer that cleaners remove ectoparasites, yet in many cases, cleaners prefer consuming mucus (Grutter and Bshary, 2003). Mucivory is typically characterized as “cheating” from the client’s point of view, often resulting in the abrupt termination of a cleaning bout (Johnstone and Bshary, 2002; Bshary and Grutter, 2005; Bshary and Grutter, 2006). While adaptations suited for mucivory may play a substantial role in shaping the morphology of these obligate cleaners, this feeding behavior has not been observed in any facultative cleaners. Thus, an analysis of the morphology of facultative cleaners may yield crucial insights about traits that are conducive to ectoparasite removal, without the complication of accounting for morphological innovations related to mucivory.

The prevalence of cleaning behavior is greatest in the mostly coral-reef-associated clade Labridae (wrasses & parrotfishes) (Coté, 2000; Froese and Pauly, 2014). The majority of cleaner wrasses, including several members of the genus *Thalassoma*, are facultative cleaners, shifting towards dietary patterns that are very similar to those of non-cleaner congeners as adults (Bellwood et al., 2006; Froese and Pauly, 2014). The approximate sizes at which these facultative cleaners stop cleaning are reported in the literature, and observations of cleaning in these species are almost exclusively restricted to individuals that display juvenile coloration patterns (Darcy et al, 1974; Losey et al., 1994; Wicksten, 1998). Rarely do facultative cleaners continue to exhibit this behavior into adulthood. A few species, such as *Thalassoma duperrey*, clean as adults, although cleaning in the adult form of this species appears to involve the removal of parasitic barnacles from turtles (Losey et al., 1994). Therefore, individuals of *Thalassoma* cleaner species that have shifted to adult colorations (regardless of sex) are far less likely to engage in ectoparasitic isopod removal. It is unknown why facultative cleaner species exhibit such shifts, or whether these species exhibit strong ontogenetic changes in morphology that accompany the change in feeding ecology.

Most members of the *Thalassoma* clade are predators of a wide variety of benthic invertebrates, including gastropods, bivalves, urchins, and crustaceans (Table S1, see supplementary materials). As many of these prey items are usually attached to a substrate, *Thalassoma* wrasses are generally expected to employ a biting mechanism to acquire prey. Biting may be especially important to prey capture in facultative cleaners in this clade, as many of these species have been described to “pick” ectoparasites off the bodies of their clients (Hobson, 1969; Hobson, 1971; Wicksten, 1998). Facultative cleaners within the *Thalassoma* clade do not constitute a monophyletic group, suggesting that cleaning behavior appears to have a dynamic evolutionary history. While we acknowledge that all taxa in our study exhibit changes in diet as they attain larger body sizes, what unites the facultative cleaner species is their engagement in cleaning behavior as juveniles. The evolution of facultative cleaning in this clade provided us with the opportunity to explore the extent to which scaling patterns in functional traits are linked to ontogenetic shifts away from cleaning behavior.

In this study, we focused on characterizing the ontogeny of functional traits in several members of the monophyletic group comprised of two genera, *Thalassoma* and *Gomphosus*. We aimed to test the hypothesis that transitions away from cleaning are associated with informative scaling patterns in the feeding apparatus. Our aims were to: 1) identify whether facultative cleaners exhibit different scaling patterns in traits related to feeding when compared to non-cleaners, and 2) to determine whether scaling patterns in feeding traits correspond with shifts away from cleaning.

Results

Allometry of traits

Disparate patterns of scaling emerged when we examined the ontogeny of vertical gape distance, maxillary KT, jaw protrusion distance, and bite force for *Thalassoma* wrasses (Fig. 1; see also

Fig. S1 in supplementary materials for species-specific patterns). The majority of *Thalassoma* wrasses exhibited some form of allometry in these functional traits. *Gomphosus varius* showed the most consistent pattern of allometry across all four traits, exhibiting negative allometry in each case. Generally, for each trait, no more than three species exhibited isometry.

All wrasses exhibited strong relationships between $\log_{10}$ vertical gape distance and $\log_{10}$ body length ($R^2 = 0.94$-0.99, all p-values < 0.0001; Table S2 in supplementary material). The SMA regression slopes ranged from 0.51 to 1.58. While non-cleaner taxa exhibited either isometry or negative allometry for vertical gape distance, facultative cleaners almost universally exhibited positive allometry (Fig. 1A). *Thalassoma duperrey* was the only facultative cleaner to exhibit isometry for this trait.

The relationships between $\log_{10}$ premaxillary protrusion distance and $\log_{10}$ body length (Fig. 1B) were strong and varied substantially ($R^2 = 0.60$-0.98, all p-values < 0.0001; Table S3 in supplementary material). The SMA regression slopes ranged from 0.62 to 2.46. Compared to an isometric slope of 1.0, seven species exhibited positive allometry; only three species (two of which were non-cleaners) exhibited isometric trends.

The relationship between $\log_{10}$ bite force and $\log_{10}$ body length (Fig. 1C) was strong across all species ($R^2 = 0.92$-0.99, all p-values < 0.0001; Table S4 in supplementary material). The SMA regression slopes ranged from 1.16 to 2.96. All but one of the facultative cleaner species (*T. bifasciatum*) showed positive allometry for this trait, while only one non-cleaner (*T. rueppellii*) shared this trend.

Wrasses in this study universally showed strong negative allometry for maxillary KT (Fig. 1D). The SMA regressions yielded slopes that ranged from -0.38 to -1.08 ($R^2 = 0.90$-0.98, all p-values < 0.0001; see Table S5 in supplementary material). All t-tests designed to test the hypothesis that slope was different from 0 (as predicted by isometry) indicated strong, significant deviations from isometry (all p-values < 0.0001).

Allometric differences between facultative cleaners and non-cleaners

Through a series of Šidák-corrected two-sample t-tests, we found significant differences in mean slopes between facultative cleaners and non-cleaners in only two traits: $\log_{10}$ vertical gape distance and $\log_{10}$ bite force (Table 1). Here, we observed that facultative cleaners exhibited steeper slopes for both $\log_{10}$ vertical gape distance ($t_9 = 4.11$, $p = 0.021$; Table 1) and $\log_{10}$ bite force ($t_9 = 2.66$, $p = 0.046$; Table 1).

Results of the Wilcox procedure

We employed the Wilcox procedure to each trait analysis to compare the regression line of each non-cleaner with those of facultative cleaners, allowing us to identify the regions where the data in each pairwise comparison begin to overlap. For $\log_{10}$ vertical gape distance, the regression line of each facultative cleaner intersected those of 1 to 4 non-cleaner species (Table 2). In 18 cases (out of 30 total comparisons), the Wilcox procedure identified regions of overlap beginning at $\log_{10}$ body lengths of 1.81 +/- 0.0081 (mean +/- s.d.), indicating that overlap of data did not occur until species attained a body length of 63.98 +/- 12.03 mm. Thus, these results indicate that juvenile facultative cleaners smaller than 63.98 +/- 12.03 mm exhibited significantly lower vertical gape distances compared to non-cleaners (Fig. 1A).

Our analysis of $\log_{10}$ premaxillary protrusion distance indicated that the regression line of each facultative cleaner intersected those of 1 to 3 non-cleaner species (Table 2). In 9 cases, the Wilcox procedure identified regions of overlap beginning at $\log_{10}$ body lengths of 1.79 +/- 0.101 (mean +/- s.d.), revealing that overlap in regression lines did not occur until pairs of species attained body lengths of 62.12 +/- 14.52 mm (Fig. 1B). The nature of regression line overlap for this trait presents a heterogeneous pattern; here, two non-cleaner species, *T. amblycephalum* and *T. quinquevittatum*, exhibited smaller premaxillary protrusion distances as juveniles than did some of the facultative cleaner species (Fig. 1B).

For $\log_{10}$ bite force, the regression line of each facultative cleaner intersected those of 1 to 4 non-cleaner species (Table 2). In 14 total cases, the Wilcox procedure revealed regions of overlap beginning at $\log_{10}$ body lengths of 1.78 +/- 0.0092 (mean +/- s.d.). This indicated that overlap in bite force did not occur until both species reached a body length of 59.50 +/- 12.71 mm (Fig. 1C).

The Wilcox procedure identified 3 cases of significant overlap in regression lines for maxillary KT (Table 2). The regression line of each facultative cleaner intersected those of 1 to 5 non-cleaner species. The regions of overlap began at $\log_{10}$ body lengths of 2.05 +/- 0.13 (s.d.), indicating that overlap did not occur until species reached body lengths of 112.05 +/- 32.65 mm (Fig. 1D).

For all of the above traits, with the exception of maxillary KT, the $\log_{10}$ body lengths at which data overlap began (identified by the Wilcox procedure) were similar. To understand whether facultative cleaners and non-cleaners converged at critical body sizes that correspond to juvenile or adult color phases, we conducted a *post-hoc* analysis. We used a two-sample *t*-test of means to compare the body lengths of the largest specimens with juvenile color patterns in facultative cleaners with those in non-cleaners. We found no significant difference in these body lengths between facultative cleaner and non-cleaner species ($t_9 = -0.78$, $p = 0.46$). Our recorded range of body lengths for specimens with juvenile coloration patterns and those with adult coloration patterns is shown in Table S6 (see supplementary materials). The body lengths of the largest juvenile-colored specimen in each facultative cleaner species was 64.28 +/- 3.12 mm (mean +/- s.d.), while the body length of the largest juvenile-colored non-cleaner specimen was 66.14 +/- 4.78 mm (mean +/- s.d.).

Investigating trends in bite force

Our multiple regression analysis captured a large amount of the variation in bite force for each species (Adjusted $R^2 = 0.93$-0.99, all *p*-values < 0.00001). In general, the $\log_{10}$ mass of the A2 or A3 muscle had the largest CAR score, revealing that these variables contribute the greatest to bite force in both facultative cleaners and non-cleaners (Table 3). In two cases (corresponding to *T. hardwicke*, a non-cleaner, and *T. pavo*, a facultative cleaner), the fiber length of the A3 had the highest CAR score. We thus found that ontogenetic changes to muscle sizes generally contributed more strongly to the ontogenetic patterns in bite force than did changes in the mechanical advantage associated with either muscle. Using MANOVA, we found no significant differences between the CAR scores of all six variables according to the category (facultative cleaner vs. non-cleaner) to which species belonged (Wilk’s $\text{Lambda}_{6,4} = 0.54$, $p = 0.74$).

The allometric patterns of each of the six traits we used in the multiple regression analyses are depicted in Fig. 2. The ontogenetic patterns of residual mechanical advantage for the A2 and the A3 muscles varied widely across species (Fig. 2A, B). Slopes varied from -0.12 to 0.16 for the former and from -0.22 to 0.25 for the latter (Tables S6 and S7, see supplementary materials).

Over ontogeny, both facultative cleaners and non-cleaners generally exhibited negative allometry or isometry for $\log_{10}$ A2 mass (Fig. 2C). Only two species, *T. lutescens* and *G. varius*, showed positive allometry. We found similar results when analyzing the ontogeny of $\log_{10}$ A3 mass (Fig. 2D). A single outlier, *T. lutescens*, exhibited the only positively allometric relationship, with a slope value of 4.10. Further SMA regression information for these traits is available in Tables S8 and S9 (see supplementary materials).

Most species (n=7) exhibited negative allometry for $\log_{10}$ A2 fiber length (Fig. 2E), with two species, *T. rueppellii* and *T. amblycephalum* characterized by extremely shallow slopes (0.46 and 0.66, respectively). An additional three species showed isometry for this trait, while only *T. duperrey* and *T. pavo* exhibited positive allometry. We observed similar patterns for $\log_{10}$ A3 fiber length. Again, *T. rueppellii* and *T. amblycephalum* showed markedly shallower slopes than most species (0.51 and 0.59, respectively), and *T. duperrey* and *T. pavo* again exhibited positive allometry. Further SMA regression information for these traits is available in Tables S10 and S11 (see supplementary materials).

A series of Šidák-corrected two-sample *t*-tests found no differences between slope means between facultative cleaners and non-cleaners for any of these bite-force-related traits (Table 1).

Discussion

Our data reveal that as juveniles, facultative cleaner fishes tended to exhibit lower bite forces, lower maxillary KTs, and smaller vertical gape distances compared to juvenile non-cleaner congeners. In a number of cases, facultative cleaners exhibited smaller premaxillary protrusion distances as juveniles, but this was not consistent across all species. There was remarkable consistency in the range of body lengths over which we observed significant differences between facultative cleaners and non-cleaners in bite force, vertical gape distance, and premaxillary protrusion distance. Facultative cleaner species generally appeared to exhibit lower trait values until they attained, on average, a body length of roughly 62 mm. This body length coincides closely with the approximate body lengths at which facultative cleaners in our dataset switched from juvenile to adult coloration (roughly 64 mm). As body length increased beyond this threshold, the disparity in functional traits between facultative cleaners and their non-cleaner relatives no longer achieved statistical significance. Thus, the body lengths at which we first found overlap between facultative and non-cleaner species in feeding traits correspond with the body lengths over which these species shift from juvenile to adult coloration patterns. Only in the case of maxillary KT did the point at which facultative cleaners and non-cleaners exhibit substantial overlap occur well after the phase transition for each pair of species.

Collectively, these regression analyses enable us to summarize the chronology of the ontogeny of feeding traits in facultative cleaners and non-cleaners. As *Thalassoma* wrasses increase in body size, disparities in bite force are the first to disappear, driven primarily by the scaling of A2 or A3 muscle masses. Increases in bite force are followed by increases in premaxillary protrusion and vertical gape distance. Facultative cleaners do not show convergence in maxillary KT with non-cleaner congeners until well into adulthood.

The ontogeny of maxillary KT in wrasses

Maxillary KT is often used to assess the functional implications of diverse shapes observed in the anterior-jaws four-bar linkage system in wrasses (Westneat, 1995; Hulsey and Wainwright, 2002; Wainwright et al., 2004). Such studies often generate species-mean values of maxillary KT to assess functional diversity across species, yet few have explored intraspecific diversity in this trait (but see Westneat, 1994). The large range of maxillary KTs exhibited by individuals in our dataset is intriguing. *Thalassoma hardwicke*, a non-cleaner, exhibited the largest range of maxillary KT for a species, with a range of 0.64 (from 1.45 to 0.81). Surprisingly, the range of maxillary KTs exhibited over the ontogeny of this species alone overlapped species-mean maxillary KT values reported for 66 species of wrasses and parrotfishes on the Great Barrier Reef (Wainwright et al., 2004). Our data support the notion that ontogenetic trajectories have the potential to increase functional disparity in an already diverse system. Though *T. hardwicke* exhibited the largest range in maxillary KT over ontogeny, the median range for species in our dataset was 0.51. Thus, *Thalassoma* wrasses exhibited substantial change in maxillary KT over ontogeny. Our findings reveal the importance of considering the ontogenetic trajectories of traits when quantifying inter- and intra-specific patterns of diversity.

The morphological basis for cleaning in Thalassoma

Behavioral descriptions of juvenile fishes continuously "picking" ectoparasites from the bodies of their respective client organisms have contributed towards our superficial understanding of cleaning behavior (Darcy et al., 1974; Losey et al., 1994). Previously, researchers have described the mouth movements of cleaner fishes as precisely and repetitively "picking" ectoparasites off clients (Hobson, 1969; Hobson, 1971; Darcy et al., 1974; Wicksten, 1998). The functional morphology of this behavior, however, has not been systematically studied, and details of how prey capture occurs are lacking. The term "picking" has also been used to describe a form of biting prey capture in some cichlids (Liem, 1979), and embiotodicids and labrids (Horn and Ferry-Graham, 2006), whereby small, sessile prey are dislodged from a substrate. Here, precise and repeated movements of the upper jaws allow protruding teeth to be used as a prehensile tool (Liem, 1979). In Cyprinodontiformes, "picking" is also used to describe precisely-controlled and coordinated ''forceps-like'' movements of the upper and lower jaws (Ferry-Graham et al., 2008; Hernandez et al., 2009). In contrast to other forms of biting, cyprinodontiform picking involves the acquisition of individual prey items (small invertebrate prey) from the substrate or water column, while other items are left behind (Hernandez et al., 2008). The fine-tuned precise movements underlying the picking behavior in cyprinidontiform taxa are associated with a morphological novelty in the premaxillomandibular ligament connecting the upper and lower jaws (Hernandez et al., 2008). Whether picking in cleaner fishes is similar to the picking behavior observed in other taxa is unknown and can only be determined with future kinematic studies. We hypothesize that cleaner fishes employ a similar feeding strategy using precise, coordinated movements of the jaws; such would allow the cleaner to remove targeted items from a client's body, leaving little room for error in haphazardly biting into the client itself. In the present study, we find that a combination of traits (a small vertical gape, small bite force, low maxillary KT, and, in some cases, a low premaxillary protrusion distance) underlie the "picking" seen in facultative *Thalassoma* cleaner fishes.

A low maxillary KT in cleaners (compared to that of juvenile non-cleaners) indicates an alteration of the four-bar linkage system that reduces the overall displacement of the structures involved. While oral jaws with higher values of maxillary KT tend to be characterized as "velocity-modified" (Westneat, 1994), this categorization may better apply to suction-based prey capture events, wherein the jaws need to rapidly expand (Westneat, 1994; Hulsey and

Wainwright, 2002; Wainwright et al., 2004). In the context of biting behaviors, it may be better to characterize systems with higher values of KT as “displacement-modified”, as the actions of jaw opening and closing involve greater overall displacement. If the “picking” behavior of cleaners occurs via biting attached prey in a repetitive, cyclical manner, an overall reduction in displacement (i.e. lower maxillary KT) could allow for a higher frequency of the bite cycle. Here, a small vertical gape and low premaxillary protrusion would similarly reduce the overall displacement incurred during a single bite cycle of opening and closing. Furthermore, as noted by Hulsey and Wainwright (2002), lower values of KT transmit more force to the maxilla as the lower jaw closes. This may further aid species that consume prey attached to a substrate by maximizing force transmission to the upper jaw when the oral jaws make contact to bite into or pull off the prey. Corroborated by the observed reduction in vertical gape, and in some cases, premaxillary protrusion distance, our results indicated that juvenile *Thalassoma* facultative cleaner fishes may rely on low-displacement, rapid bite cycles to consume ectoparasites attached to their client species.

Significant differences in bite force between facultative cleaners and non-cleaners in the juvenile condition add to the functional disparities between these two groups. Facultative cleaners are limited in bite force as juveniles, which appears to present little problem to cleaning, given that the ectoparasites consumed have thus far been reported to be small (typically <1.5 mm diameter), thin-shelled juvenile gnathiid or cymothoid isopods (Losey, 1972; Losey, 1974; Darcy et al., 1974; Davies, 1981; Grutter, 1996). Through ontogeny, however, the strong positively-allometric trends in bite force observed in these species enable them to exhibit bite forces similar to those of adult non-cleaner congeners, potentially allowing facultative cleaner species to exploit new prey or expand their prey breadth. Though this allometry seems to be largely driven by ontogenetic changes to the masses of the A2 and A3 subdivisions, this pattern of development was not unique to facultative cleaners. We were unable to find a clear, consistent pattern to indicate the mechanism via which facultative cleaner species undergo stronger allometry in bite force than do their non-cleaner congeners. It thus appears that facultative cleaner fishes experience sharp allometric trends in bite force through a diversity of ontogenetic patterns.

Why clean?

Juvenile facultative cleaners ostensibly appear to be at a functional disadvantage to sympatric, juvenile non-cleaner congeners. We suggest that these reduced suites of mechanical features, low maxillary KT, small gape, and low bite force, in *Thalassoma* wrasses may have evolved as a result of competitive displacement. *Thalassoma* wrasses co-occur, often in great abundances, in many tropical regions including but not limited to the Indian Ocean, the Indo-Pacific, and the Caribbean (Randall et al., 1997; Geange, 2010; Geange et al., 2013; Froese and Pauly, 2014). In fact, field surveys conducted in the Indo-Pacific and Caribbean found that reef crests were typically numerically dominated by members of this clade (Bellwood et al. 2002). Since each facultative cleaner wrasse in our study overlaps in habitat with other *Thalassoma* (in addition to other wrasses), juveniles of each of these species compete for prey directly with saturated assemblages of adult conspecifics and congeners. Geange et al. (2013) show that even among juvenile non-cleaner taxa in this clade, intense asymmetric competition between congeners can result within areas of high densities. Our results suggest that cleaning behavior may present a mechanism by which such competition is reduced by providing juvenile facultative cleaner species an alternate source of prey. Through ontogeny, these facultative cleaner fishes exhibit an allometric increase in vertical gape distance, bite force, and a heavier reliance on the forceful movement of the jaws, which together possibly obviate the need to continue cleaning.

Methods

Specimen acquisition

We collected an ontogenetic series for 11 wrasse species, all members of the *Thalassoma* clade (Fig. 3). We borrowed specimens from museums (see Table S6 for accession numbers, see supplementary materials), but for two species, *Gomphosus varius* and *Thalassoma bifasciatum*, we purchased additional live specimens from fish wholesalers to assemble their complete ontogenetic series. We euthanized live specimens via an overdose of MS-222 (IACUC protocol 1006) and fixed them in 10% buffered formalin for 10-14 days before transferring them to 70% ethanol for short-term storage.

Each ontogenetic series included a size range of individuals that encompassed small juvenile sizes (~40 mm standard length) through the adult common length (or maximum length) reported for the species on FishBase (Froese and Pauly, 2014). We recorded whether each specimen exhibited a juvenile or an adult color pattern (regardless of sexual dimorphism in color) by referencing specimens with guidebooks (Burgess et al., 1991; Randall et al., 1997; Myers, 1999). We performed an extensive literature search to gather dietary information for each species and to determine which species are known to clean as juveniles. We found that across the species, the adult diet is generally similar and includes a variety of benthic crustaceans. Only in the case of *T. amblycephalum* did we observe a distinct feeding strategy: zooplanktivory. Information for all species is available in Table S1 (see supplementary material).

Using body length to determine the effects of size

To assess how morphological and functional traits change with increasing body size, we adopted a metric of size that was applicable to all species. Because of the extreme elongation of the jaws in *G. varius*, a trait which itself changes over ontogeny (Myers, 1999), standard length was not a universally applicable metric of size. Instead, we computed ‘body length’ by subtracting head length (defined here as the linear distance from the tip of the snout to the posterior edge of the operculum) from standard length. We took measurements to the nearest 0.01 mm using digital calipers.

Bite force estimation

We dissected each section of the adductor mandibulae from one side of each specimen. In wrasses, the adductor mandibulae complex originates broadly from the lateral surface of the suspensorium and inserts via tendons on the maxilla and articular bones (Winterbottom, 1974). This muscle complex is responsible for generating the force that powers the closing of the jaws during biting. Jaw closing in this muscle complex is accomplished via three subdivisions: A1, A2, and A3 (Fig. 4A). The A1 subdivision, which inserts on the maxilla, functions as part of the upper jaw protrusion mechanism. Following Westneat (2003), we did not include the A1 in our bite force model to focus on the major jaw closers: A2 and A3. Although we removed each of the three subdivisions, we only weighed the A2 and A3 subdivisions to the nearest 0.001 g using a Secura 213-1S precision balance (Sartorius Stedim Biotech GmbH, Germany).

Following muscle dissections, we cleared and double-stained specimens for bone and cartilage following Dingerkus and Uhler (1977). We then took 12 linear measurements of the skull for use with the application MandibLever 3.3 (Westneat, 2003); see Fig. 4B for measurement definitions. We collected all linear measurements using an ocular micrometer attached to a Leica M80 light microscope (Leica Microsystems GmbH, Wetzlar, Germany). MandibLever 3.3 creates a model of the lower jaw of fishes and uses musculoskeletal data to predict force transmission across the system. We used MandibLever 3.3 to estimate the total bite force (in newtons) for each specimen in this study, thus enabling us to compare the ontogenetic transitions in bite force across species.

Kinematic transmission coefficient (KT) and morphological traits

In wrasses, a four-bar linkage system in the anterior jaws guides the rotation of the maxilla and the protrusion of the premaxilla as the mandible is depressed (Westneat, 1990). The kinematic transmission coefficient (KT) for this linkage system relates the amount of maxillary rotation produced by a given amount of lower jaw rotation. This ratio is analogous to the inverse of the mechanical advantage of simple lever systems. We characterized the linkage mechanics of the anterior-jaw four-bar linkage system (Fig. 4C-D) following methods used by Wainwright et al. (2004). As shown in Figs. 4C-D, the fixed link (f) is defined as the distance between the quadrate-articular joint and the proximal base of the nasal bone. The coupler link (c) is the distance from the proximal base of the nasal to the distal end of the nasal. The output link (o) is the distance from the distal end of the nasal to the confluence between the distal end of the alveolar arm of the premaxilla, the distal arm of the maxilla and the coronoid process of the mandible. The input link (i) is thus the distance between the latter point and the quadrate–articular joint. Depression of the mandible and rotation of the maxilla protrude the premaxilla (Fig. 4D).

We calculated the maxillary KT as the ratio between the degrees of maxillary rotation and the degrees of lower jaw rotation for each specimen (Westneat, 1990). Because of the nature of this ratio, maxillary KT is dimensionless. To maintain consistency across all specimens, we measured all starting angles with the jaws closed. We then rotated the lower jaw into a fully depressed position to quantify the changes in the angles associated with the input and output to the four-bar system.

In addition to maxillary KT, we measured jaw protrusion distance and vertical gape distance. Jaw protrusion distance is the excursion distance of the most anterior canine tooth on the premaxilla as the upper jaw travels rostrally when the lower jaw is depressed. Vertical gape distance is the distance between each of the most anterior canine teeth on the upper and lower jaws when the mouth was fully open. For each of these measurements, we rotated the lower jaw into a fully depressed position without forcing it beyond natural extension. Measurements were recorded to the nearest 0.01 mm using the program ImageJ 1.47 (Rasband, 2014).

Allometry of traits

We performed all statistical analyses in R 3.0.0 (R Development Core Team, 2013). To determine the scaling pattern of traits for each species, we performed a Standardized Major Axis (SMA) regression between the $\log_{10}$ trait value (except in the case of maxillary KT, for which we used raw values) and $\log_{10}$ body length using the *smatr* package (Warton et al., 2012) in R. We then compared these regression slopes with null predictions based on isometric scaling (for ratios = 0.0; linear distances = 1.0; areas and forces = 2.0). Here, we used modified *t*-tests to determine whether the observed slopes deviated significantly from the null predictions, which are based on Euclidean geometry (Hill, 1950; Schmidt-Nielson, 1984; Emerson and Bramble, 1993; Pfaller et al., 2011). We considered deviations from isometry to be significant if the predicted slopes fell outside the 95% confidence intervals of the slopes we observed. We interpreted positive or negative deviations from isometry in these regression slopes as positive or negative allometry, respectively.

Lastly, we analyzed the ontogeny of the mechanical advantage of each muscle subdivision-specific lever system. To avoid the bias of such ratios in analysis (Packard and Boardman, 1988), we computed the residuals from a linear regression of each $\log_{10}$ in-lever (dependent variable) against the $\log_{10}$ out-lever (covariate) in an ANalysis of COVAriance (ANCOVA), using species as the independent variable. We then used each set of residuals in subsequent analyses. These are hereafter referred to as the residual mechanical advantage of the A2 and the A3, respectively.

Testing for phylogenetic signal

For each trait, we used the SMA regression slope between the $\log_{10}$ trait value (except in the case of maxillary KT, for which we used raw values) and $\log_{10}$ body length (outlined above). We then tested for phylogenetic signal in the slopes, treating the slope of each regression as its respective species' trait value. We estimated Pagel's lambda (Pagel, 1999) and Blomberg's K (Blomberg et al., 2003) using the package *phytools* (Revell, 2012) in R, and we tested the hypothesis that the phylogenetic signal was greater than 0. The phylogenetic tree we used, which contained all species in this study, was pruned from Kazancioglu et al. (2009), who used a supermatrix approach to propose relationships between 252 labrid species. For all traits, we found the level of phylogenetic signal to be extremely low, and we were unable to reject the null hypothesis that phylogenetic signal equals 0 (Table S13, see supplementary materials). This implies that the ontogenetic trajectories in the traits we analyzed have been under such strong selection, that the trends in slopes do not follow a Brownian motion model of evolution. While we cannot rule out that the lack of phylogenetic signal may be due to limited power, we were unable to detect phylogenetic signal in any of the traits we analyzed. Therefore, all subsequent analyses were performed using traditional parametric statistical methods.

Comparisons between facultative cleaners and non-cleaners

To evaluate the relationship between each trait and body length across groups of species, we first tested for homogeneity of slopes by building separate General Linear Models (GLMs), specific to each trait. In each case, the dependent variable was the $\log_{10}$ trait value (except in the case of maxillary KT, for which we used raw values), species identity was the independent variable, and $\log_{10}$ body length was used as a covariate. We tested the hypothesis that interactions between the independent variable (species) and the covariates ($\log_{10}$ body length) were significantly different from 0. In every case, we found highly significant interactions, indicating that in each analysis, the species-specific slopes were not homogeneous. As this violates a key assumption of ANCOVA, we refrained from comparing least squares means across species.

To test the hypothesis that facultative cleaners exhibit more extreme allometry in functional traits compared to non-cleaners, we performed a series of trait-specific two-sample t-tests. For each trait, we assessed the equality of the mean of the SMA regression slopes for the facultative cleaners against that of the non-cleaners. Here, the null hypothesis was that the mean SMA regression slope would not differ between the two groups. We applied a Šidák correction to reduce the Type I error probability across multiple comparisons (Šidák, 1967).

To identify the morphological correlates of cleaning, we made comparisons of trait magnitudes between facultative cleaners and non-cleaners across the range of body lengths. Given the heterogeneity of slopes, we employed the Wilcox procedure (Wilcox, 1987) to determine regions of the x-axis ($\log_{10}$ body length in all cases) in which trait values for non-cleaners were significantly different from those of facultative cleaners. This allowed us to distinguish at what body lengths the differences in trait magnitudes seen between facultative and non-cleaner species were no longer significantly different, given the error structures of the regressions. We decided to adopt this approach rather than use *t*-tests to compare regression intercepts because such intercepts represent trait values at a body size of zero, and thus do not constitute trait values that are biologically realistic. The Wilcox procedure is a modification of the Johnson-Neyman method (Johnson and Neyman, 1936) that is adjusted for multiple comparisons. For each trait, we compared the regression line of every non-cleaner species (n=6) with that of each facultative cleaner species (n=5), for a total of 30 comparisons per trait. The Wilcox procedure allowed us to identify the regions where the data in each pairwise comparison begin to overlap, taking into account the spread of data around each regression line. In several cases, regression lines crossed at values of x that represented biologically-impossible body lengths for either or both of the species involved. We therefore restricted values of x to those that were covered within our dataset.

Investigating trends in bite force

To understand the factors that determine the ontogeny of bite force in each species, we ran separate, species-specific multiple regression analyses. In each analysis, we used the $\log_{10}$ bite force as calculated by MandibLever 3.3 as the dependent variable. For the independent variables, we used: the residual mechanical advantage values of the A2 and A3, the $\log_{10}$ fiber lengths of the A2 and A3, and the $\log_{10}$ masses of the A2 and A3. We selected these variables because MandibLever 3.3 computes muscle cross-sectional area as muscle mass divided by muscle fiber length times muscle density (Westneat, 2003), and assumes a constant muscle density for all specimens (McMahon, 1984). Our measurement of muscle fiber length follows that of Westneat (2003). Both mass and fiber length were included as predictors because force generated by a muscle is proportional to the muscle's cross-sectional area (Powell et al., 1984).

We then used the R package *relaimpo* (Grömping, 2006) to calculate the R^2 decomposition of each model according to Zuber and Strimmer (2010). This method computes Correlation-Adjusted coRrelation (CAR) scores, which are the correlations between the response and the Mahalanobis-decorrelated predictors. Comparing these scores allowed us to assess the relative importance of each predictor, and enabled us to identify which variable most strongly predicted each species' ontogenetic bite force trajectory. We then used Multiple Analysis Of VAriance (MANOVA) to assess whether facultative cleaners exhibited significant differences in the CAR scores of all six variables when compared to non-cleaners.

List of Symbols and Abbreviations

Museums

CAS	California Academy of Sciences
LACM	Los Angeles County Museum of Natural History
USNM	Smithsonian National Museum of Natural History
VB	Vikram Baliga's personal collection

Mechanics

KT	kinematic transmission coefficient
f	fixed link, in anterior-jaw four-bar linkage system
c	coupler link, in anterior-jaw four-bar linkage system
i	input link, in anterior-jaw four-bar linkage system
o	output link, in anterior-jaw four-bar linkage system

Statistics

s.d.	standard deviation

R^2	coefficient of determination
SMA	standardized major axis regression
ANCOVA	analysis of covariance
MANOVA	multiple analysis of variance
t	t-value; subscript denotes degrees of freedom

Allometry (in Appendix only)

I	isometry
-	negative allometry
+	positive allometry

Acknowledgements

We would like to thank collections managers Jeffrey T. Williams (USNM), Rick Feeney (LACM), and David Catania (CAS) for their invaluable help in approving and processing specimen loans. We thank Benjamin A. Higgins, Christopher J. Law, and Sarah Kienle for valuable comments on the manuscript. Christopher J. Law prepared the illustrations shown in Fig. 1 from photographs of specimens. David C. Collar and Peter T. Raimondi provided critical advice towards our statistical analyses. Finally, we thank undergraduate research assistants Michellé Mac, Michaela Tondi, Brandon Parks, Kathryn Pelon, and Danielle Pruitt for their assistance in processing and clearing and staining specimens.

Funding

This work was supported by the Society for the Study of Evolution’s Rosemary Grant Award for Graduate Student Research, awarded to V.B.B.

Author Contributions

V.B.B. collected all specimens and data, performed all data analyses, and prepared the manuscript. R.S.M. helped develop the approach, provided crucial insights, and edited the manuscript.

Author Competing Interests

The authors have no competing interests to declare.

References

Bellwood, D. R., Wainwright, P. C., Fulton, C. J., and Hoey, A. S. (2006). Functional versatility supports coral reef biodiversity. *Proc. R. Soc. B.* **273**(1582), 101–107.

Blomberg, S. P., Garland Jr., T., and Ives, A. R. (2003). Testing for phylogenetic signal in comparative data: Behavioral traits are more labile. *Evolution* **57**, 717 -745.

Bronstein, J. L. (1994). Conditional outcomes in mutualistic interactions. *Trends Ecol. Evolut.* **9**(6), 214–217.

Bshary, R., and Grutter, A. S., (2002). Asymmetric cheating opportunities and partner control in the cleaner fish mutualism. *Animal Behaviour*. **63**, 547-555**.**

Bshary, R., and Grutter, A. S., (2005). Punishment and partner switching cause cooperative behaviour in a cleaning mutualism. *Bio. Lett.* **1,** 396-399.

Bshary, R., and Grutter, A. S., (2006). Image scoring and cooperation in a cleaner fish mutualism. *Nature* **441,** 975-978.

Burgess, W. E., Axelrod, H. R., and Hunziker, R. (1991). *Atlas of Marine Aquarium Fishes.* Neptune City, NJ: T.F.H. Publications, Inc.

Coté, I.M. (2000). Evolution and Ecology of Cleaning Symbioses in the Sea. *Oceanograph. Mar Biol: Ann. Rev.* **38**, 311–355.

Darcy G. H., Maisel E., and Ogden T. C. (1974). Cleaning preferences of the gobies *Gobiossoma evelynae* and *G. prochilos* and the juvenile wrasse *Thalassoma bifasciatum*. *Copeia* **1974**, 375–379.

Davies, A. J. (1981) A scanning electron microscope study of the praniza larva of *Gnathia maxillaris* Montagu (Crustacea, Isopoda, Gnathiidae), with special reference to the mouthparts. *J. Natl. Hist*. **15**, 545-551

Deban, S.M. and O'Reilly, J.C. (2005). The ontogeny of feeding kinematics in a giant salamander *Cryptobranchus alleganiensis*: Does current function or phylogenetic relatedness predict the scaling patterns of movement? *Zoology* **108**,155-167.

Dingerkus G, Uhler LD. (1977). Enzyme clearing of alcian blue stained whole small vertebrates for demonstration of cartilage. *Stain Technol*. **52**, 229–232.

Emerson, S. B. and Bramble, D. M. (1993). Scaling, allometry, and skull design. In *The Skull, Vol. 3, Functional and Evolutionary Mechanisms* (ed. J. Hanken and B. K.Hall), pp. 384-421. Chicago: University of Chicago Press.

Ferry-Graham, L. A., Gibb, A. C., and Hernandez, L. P. (2008). Premaxillary movements in cyprinodontiform fishes: An unusual protrusion mechanism facilitates "picking" prey capture. *Zoology* **111**, 455-466.

Frédérich, B., and Sheets, H.D. (2009). Evolution of ontogenetic allometry shaping giant species: a case study from the damselfish genus *Daschyllus* (Pomacentridae). *Bio. J. Linn. Soc. Lond*. **99**, 99-117.

Froese, R. and Pauly, D. Editors. (2014). FishBase. World Wide Web electronic publication. www.fishbase.org, version (1/2014).

Geange, S.W. (2010). Effects of larger heterospecifics and structural refuge on the survival of a coral reef fish, *Thalassoma hardwicke*. *Mar. Ecol. Prog. Ser.* **407**, 197–207

Geange, S.W., Stier, A.C., Shima, J.S. (2013). Competitive hierarchies among three species of juvenile coral reef fishes. *Mar. Ecol. Prog. Ser.* **472**, 239–248

Gorlick, D. L. (1980). Ingestion of host fish surface mucus by the Hawaiian cleaning wrasse, *Labroides phthirophagus* (Labridae), and its effect on host species preference. *Copeia* **1980**, 863–868

Grömping, U. (2006). Relative Importance for Linear Regression in R: The Package *relaimpo*. *J. Stat. Soft.* **17**(1), 1-27.

Grutter, A. S. (1996). Parasite removal rates by the cleaner wrasse *Labroides dimidiatus*. *Mar. Ecol. Prog. Ser.* **130**, 61-70.

Grutter, A. S. (1997). Size-selective predation by the cleaner fish *Labroides dimidiatus*. *J. Fish Bio.* **50,** 1303-1308.

Grutter, A. S., and Bshary, R. (2003). Cleaner fish prefer client mucus: support for partner control mechanisms in cleaning interactions. *Proc. R. Soc. B., Bio. Letters.* **270,** Supp. 2. S242-S244.

Grutter, A. S., and Bshary, R. (2004). Cleaner fish *Labroides dimidiatus* diet preferences for different types of mucus and parasitic gnathiid isopods. *Animal Behaviour*. **68**, 583-588.

Hernandez, L. P., Ferry-Graham, L. A., and Gibb, A. C. (2008). Morphology of a picky eater: A novel mechanism underlies premaxillary protrusion and retraction within cyprinodontiforms. *Zoology* **111**, 442-454.

Hernandez, L. P., Gibb, A. C., and Ferry-Graham, L. A. (2009). Trophic Apparatus in Cyprinodontform Fishes: Functional Specializations for Picking and Scraping Behaviors. *J. Morph.* **270**, 645-661.

Herrel, A. and Gibb, A. C. (2006). Ontogeny of performance in vertebrates. *Physiol. Biochem. Zool.* **79**, 1-6.

Herrel, A. and O'Reilly, J. C. (2006). Ontogenetic scaling of bite force in lizards and turtles. *Physiol. Biochem. Zool.* **79**, 31-42.

Hill, A. V. (1950). The dimensions of animals and muscular dynamics. *Sci. Prog.* **38**, 209-230.

Hobson, E. S. (1969). Comments on Certain Recent Generalizations Regarding Cleaning Symbioses in Fishes. *Pacific Science* **23**, 35-39.

Hobson, E. S. (1971). Cleaning symbioses among California inshore fishes. *Fish. Bull.* **69**(3), 491-523.

Horn, M. H., and Ferry-Graham, L.A. (2006). Feeding Mechanisms and Trophic Interactions. In *The Ecology of Marine Fishes: California and Adjacent Waters* (eds. L.G. Allen, D. J. Pondella, and M.H. Horn), pp. 397-410. Berkeley: University of California Press.

Hulsey, C. D., and Wainwright, P. C. (2002). Projecting mechanics into morphospace: disparity in the feeding system of labrid fishes. *Proc. R. Soc. B.* **269**(1488), 317–326.

Johnson, P. O., and Neyman, J. (1936). Tests of certain linear hypotheses and their application to some educational problems. In *Statistical Research Memoirs* (eds. J. Neyman and E.S. Pearson), **1**, 57-93.

Johnstone, R. A., and Bshary, R. (2002). From parasitism to mutualism: partner control in asymmetric interactions. *Ecol. Lett.* **5**, 634–639.

Kazancioglu, E., Near, T. J., Hanel, R., and Wainwright, P. C. (2009). Influence of sexual selection and feeding functional morphology on diversification rate of parrotfishes (Scaridae). *Proc. R. Soc. B.* **276**(1672), 3439–3446.

Leim, K. F. (1979). Modulatory multiplicity in the feeding mechanism in cichlid fishes, as exemplified by the invertebrate pickers of Lake Tanganyika. *J. Zool., Lond.* **189**, 93-125**.**

Losey, G. S. (1972). The Ecological Importance of Cleaning Symbiosis. *Copeia* **1972**(4), 820–833.

Losey, G. S. (1974). Cleaning Symbiosis in Puerto Rico with Comparison to the Tropical Pacific *Copeia* **1974**(4), 960–970.

Losey, G. S., Balazs, G. H., and Privitera, L. A. (1994). Cleaning Symbiosis between the Wrasse, *Thalassoma duperrey*, and the Green Turtle, *Chelonia mydas*. *Copeia* (**3**), 684–690.

McMahon, T. A. (1984). *Muscles, Reflexes, and Locomotion*. Princeton: Princeton University Press.

Mitteroecker, P., Gunz, P., Bernhard, M., Shaefer, K., and Bookstein, F. L. (2004). Comparison of cranial ontogenetic trajectories among great apes and humans. *J. Hum. Evol.* **46**(2004), 679-698.

Myers, R. F., (1999). *Micronesian reef fishes: a comprehensive guide to the coral reef fishes of Micronesia*, 3rd revised and expanded edition. Coral Graphics, Barrigada, Guam.

Packard, G. C. and Boardman, **T. J.** (1988). The misuse of ratios, indices, and percentages in ecophysiological research. *Physiol. Zool*. **61,** 1–9.

Pagel, M. (1999). Inferring the historical patterns of biological evolution. *Nature* **401**(6756), 877-84.

Pfaller, J. B., Gignac, P. M., and Erickson, G. M. (2011). Ontogenetic changes in jaw-muscle architecture facilitate durophagy in the turtle *Sternotherus minor*. *J. Exp. Biol.* **214**(Pt 10), 1655–1667.

Powell, P. L., Roy, R. R., Kanim, P., Bello, M., Edgerton, V. R. (1984). Predictability of skeletal muscle tension from architectural determinations in guinea pig hind limbs. *J. Appl. Physiol.* **57,** 1715-1721

Poulin, R. and Grutter, A. S., (1996). Proximate Symbioses□: Cleaning and Adaptive Explanation. *BioScience* **46**(7), 512–517.

R Development Core Team. (2013). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org.

Randall, J. E., Allen, G. R. and Steene, R. C. (1997). *Fishes of the Great Barrier Reef and Coral Sea*. Crawford House Publishing, Bathurst, Australia.

Rasband, W. S., (2014). ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, http://imagej.nih.gov/ij/, 1997-2014.

Richard, B.A. and Wainwright, P.C. (1995). Scaling the feeding mechanism of largemouth bass (*Micropterus salmoides*): kinematics of prey capture. *J. Exp. Biol*., **198**, 419–433.

Revell, L. J. (2012). phytools: An R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217-223.

Schmidt-Nielson, K. (1984). *Scaling: Why is Animal Size so Important?* Cambridge: Cambridge University Press.

Šidák, Z. K. (1967). "Rectangular Confidence Regions for the Means of Multivariate Normal Distributions". *J. Am. Stat. Assoc.* **62**(318), 626–633.

Soares, M. C., Bshary, R., and Coté, I. M., (2009). Cleaning in pairs enhances honest in male cleaning gobies. *Behavioral Ecol.* **20**, 1343-1347.

Soares, M. C., Coté, I. M., Cardoso, S. C., Oliveria, R. F., and Bshary, R. (2010). Carribbean cleaning gobies prefer client ectoparasites over mucus. *Ethology*, **116**, 1244-1248.

Wainwright, P. C., Bellwood, D. R., Westneat, M. W., Grubich, J. R., and Hoey, A. S. (2004). A functional morphospace for the skull of labrid fishes: patterns of diversity in a complex biomechanical system. *Bio. J. Linn. Soc. Lond.* **82**(1), 1–25.

Warton D. I., Duursma R. A., Falster D. S., and Taskinen, S. (2012). smatr - an R package for estimation and inference about allometric lines. *Methods Ecol. Evol.* **3**, 257–259.

Westneat, M. W. (1990). Feeding mechanics of teleost fishes (Labridae; Perciformes): A test of four-bar linkage models. *J. Morph.* **205**(3), 269–295.

Westneat, M. W. (1994). Transmission of force and velocity in the feeding mechanisms of labrid fishes (Teleostei, Perciformes). *Zoomorphol.* **114**(2), 103–118.

Westneat, M. W. (1995). Feeding, Function, and Phylogeny: Analysis of Historical Biomechanics in Labrid Fishes Using Comparative Methods. *Syst. Biol.* **44**(3), 361-383.

Westneat, M. W. (2003). A biomechanical model for analysis of muscle force, power output and lower jaw motion in fishes. *J. Theor. Biol.* **223**(3), 269–281.

Wicksten, M. K. (1998). Behaviour of cleaners and their client fishes at Bonaire, Netherlands Antilles. *J. Nat. Hist.* **32**(1), 13–30.

Wilcox, R. R. (1987). *New Statistical Procedures for the Social Sciences: Modern Solutions to Basic Problems.* Hillsdale, New Jersey: Erlbaum.

Wilson, L. A. B., and Sanchez-Villagra, M. R. (2010). Diversity trends and their ontogenetic basis: an exploration of allometric disparity in rodents. *Proc. R. Soc. B.* **277**, 1227-1234.

Winterbottom, R., (1974). A descriptive synonymy of the striated muscles of the Teleostei. *Proc. Natl. Acad. Sci. Phil.* **125**, 225–317.

Zuber, V. and Strimmer, K. (2011). "High-Dimensional Regression and Variable Selection Using CAR Scores," *Stat. App. Gen. Mol. Biol.* **10**(1), 34.

Figure Captions

Fig. 1. Scaling of functional traits in all 11 species. The following traits are shown plotted against $\log_{10}$ Body Length: (A) $\log_{10}$ Vertical Gape Distance, (B) $\log_{10}$ Premaxillary Protrusion Distance, (C) $\log_{10}$ Total Bite Force (as estimated by MandibLever 3.3), and (D) Maxillary KT. Solid, orange lines indicate regressions for facultative cleaner fishes; blue dashed lines indicate regressions for non-cleaner fishes. Vertical grey lines indicate mean and standard deviation of critical x-values indicated by the Wilcox procedure for all comparisons.

Fig. 2. Scaling of muscle sizes and mechanical advantage in all 11 species. The following traits are shown plotted against $\log_{10}$ Body Length: (A) Residual Mechanical Advantage of the A2 Muscle, (B) Residual Mechanical Advantage of the A3 Muscle, (C) $\log_{10}$ A2 Mass, (D) $\log_{10}$ A3 Mass, (E) $\log_{10}$ A2 Fiber Length, and (F) $\log_{10}$ A3 Fiber Length. Solid, orange lines indicate regressions for facultative cleaner fishes; blue dashed lines indicate regressions for non-cleaner fishes.

Fig. 3. Relationships between the 11 wrasses analyzed in this study. Highlighted species are reported to be facultative cleaner fishes as juveniles (Table S1, see supplementary materials). The phylogeny is taken from Kazancioglu et al. 2009, who used a supermatrix approach for 252 species of labrids. The tree is pruned to include only the species in this study. The illustration for each species depicts juvenile female morphology and coloration pattern.

Fig. 4. Cranial morphology of *Thalassoma lucasanum*. (A) Three subdivisions of the adductor mandibulae complex. The A1 has been detached from its origin on the preopercle and moved in order to show the position of the A3, which is medial to the A1. (B) Morphometric measurements used in MandibLever 3.3. All 12 measurement definitions and line colors follow Westneat (2003): 1, 2, and 3 are the inlever lengths of A2, A3, and jaw opening, respectively; 4 is the outlever length; 5 is the length of the A2; 6 is the length of the A3; 7 is the A3 tendon length; 8 and 9 are distances from each muscle subdivision’s origin to the jaw joint; 10 is the distance between the A2 and the A3 insertion points; 11 is the distance from the tip of the jaw to the A2 subdivision insertion; 12 is the length between the tip of opening inlever to the tip of jaw. (C-D) Four-bar linkage system in closed and open positions, respectively. *Abbreviations:* art,

articular; dt, dentary; mx, maxilla; pmx, premaxilla; pop, preopercle; f, fixed link; c, coupler link; o, output link, i, input link. All scale bars are 5 mm.

Table Captions

Table 1: Comparison of the mean slopes between facultative cleaners and non-cleaners in feeding traits. The trait listed in each row indicates the dependent variable used in species-specific SMA regressions against $\log_{10}$ Body Length. The regression slopes were then pooled according to dietary category. *P*-values are adjusted according to a Šidák correction. Bold rows indicate significant differences in mean slopes between facultative cleaners and non-cleaners.

Table 2: Determining regression line overlap between facultative cleaners and non-cleaners. The cells listed under each trait display the $\log_{10}$ body length at which the regression lines of the listed facultative cleaner and non-cleaner species began to overlap. Cells without values (indicated by the dashes) indicated cases where regression lines crossed at values that represented biologically-impossible body lengths for either or both of the species involved.

Table 3: Multiple regression analyses reveal the traits that contribute the most to bite force. [a] Correlation-Adjusted coRrelation (CAR) scores, following Zuber and Strimmer (2010). Bold values indicate the trait with the largest CAR score. [b] Calculated from a linear regression of each $\log_{10}$ in-lever (dependent variable) against the $\log_{10}$ out-lever (covariate) in an analysis of covariance (ANCOVA), using species as the independent variable.

Table 1: Comparison of the mean slopes between facultative cleaners and non-cleaners in feeding traits

Trait	Mean Difference	*t*-value	d.f.	*p*-value
$\log_{10}$ Vertical Gape Distance	**0.54**	**4.11**	**9**	**0.010**
$\log_{10}$ Premaxillary Protrusion Distance	0.43	1.39	9	0.59
$\log_{10}$ Bite Force	**0.68**	**2.66**	**9**	**0.046**
Maxillary KT	0.16	1.34	9	0.62
Residual Mechanical Advantage of A2 Muscle	0.26	1.10	9	0.94
Residual Mechanical Advantage of A3 Muscle	0.42	2.05	9	0.44
$\log_{10}$ A2 Mass	0.38	1.10	9	0.88
$\log_{10}$ A3 Mass	0.63	2.03	9	0.36
$\log_{10}$ A2 Fiber Length	0.29	2.57	9	0.17
$\log_{10}$ A3 Fiber Length	0.22	1.91	9	0.43

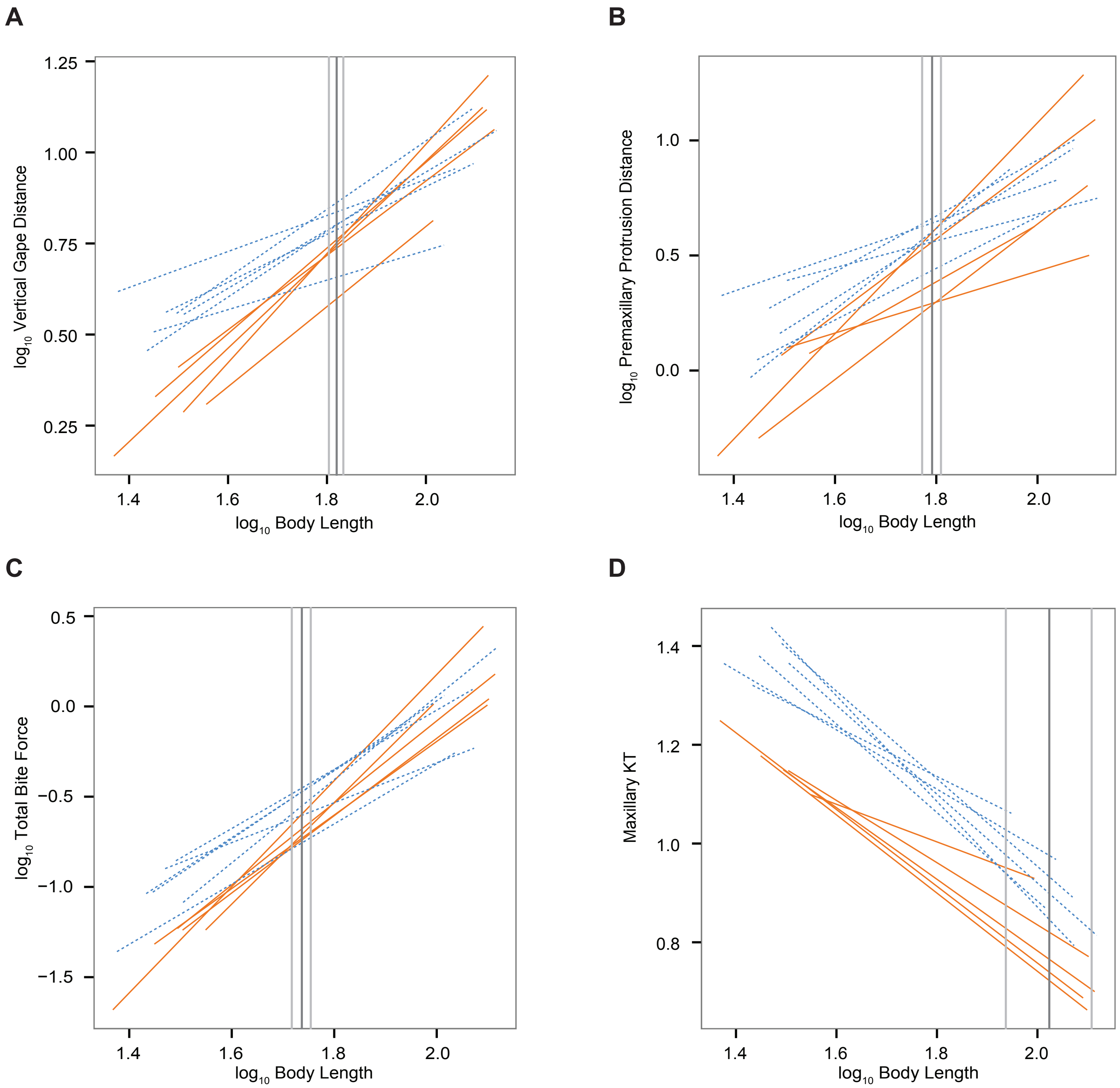
Table 2: Determining regression line overlap between facultative cleaners and non-cleaners

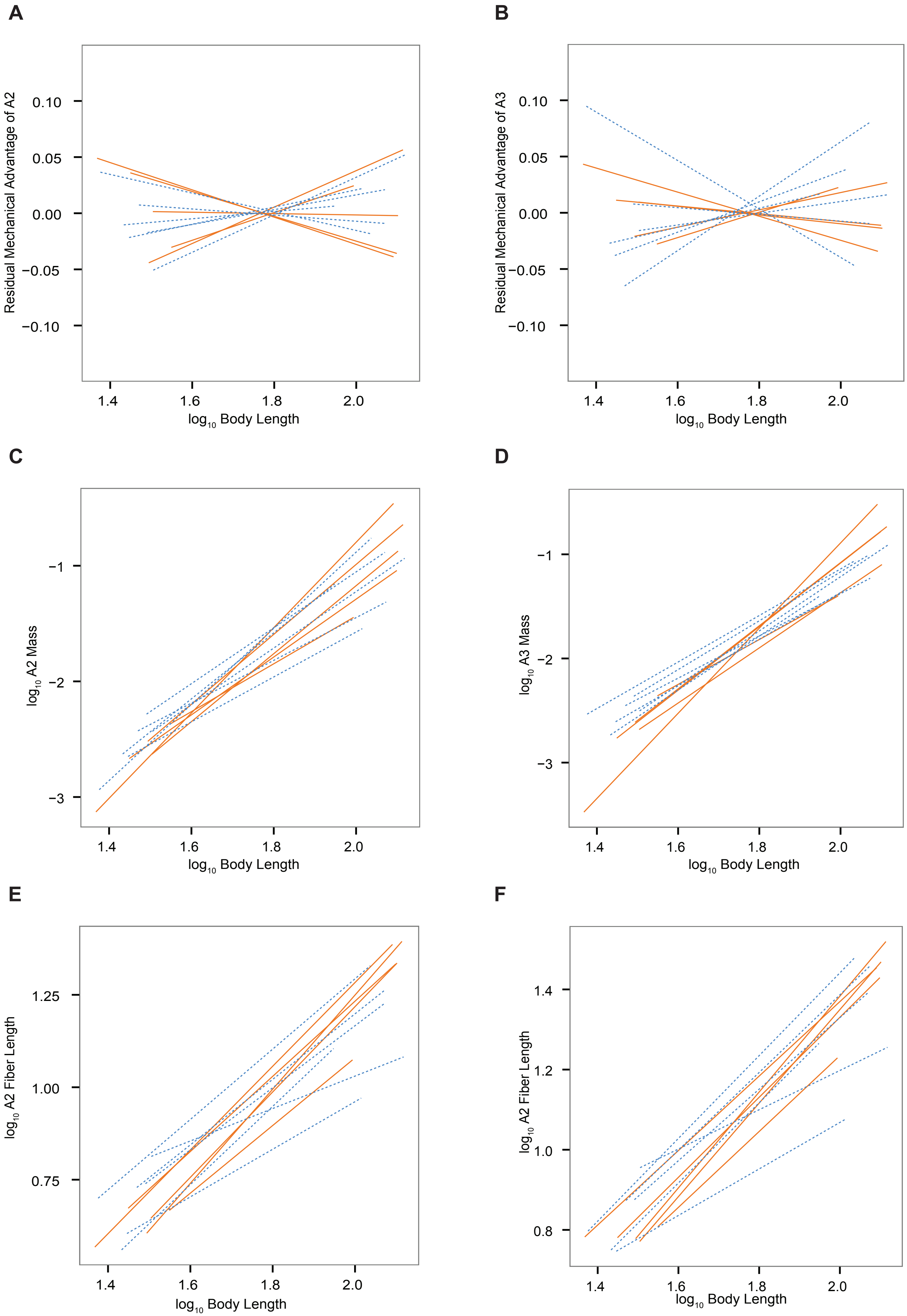
Facultative Cleaner	Non-Cleaner	$\log_{10}$ Bite Force	Maxillary KT	$\log_{10}$ Vertical Gape Distance	$\log_{10}$ Premaxillary Protrusion Distance
T. bifasciatum	*G. varius*	--	--	1.88	--
	T. amblycephalum	--	--	1.64	1.91
	T. hardwicke	1.82	2.25	1.81	--
	T. hebraicum	--	--	--	--
	T. quinquevittatum	--	--	1.83	--
	T. rueppellii	--	--	1.84	--
T. duperrey	*G. varius*	--	--	--	1.80
	T. amblycephalum	--	2.14	1.65	--
	T. hardwicke	1.76	2.11	1.88	1.86
	T. hebraicum	--	--	--	--
	T. quinquevittatum	--	--	--	--
	T. rueppellii	--	--	--	1.72

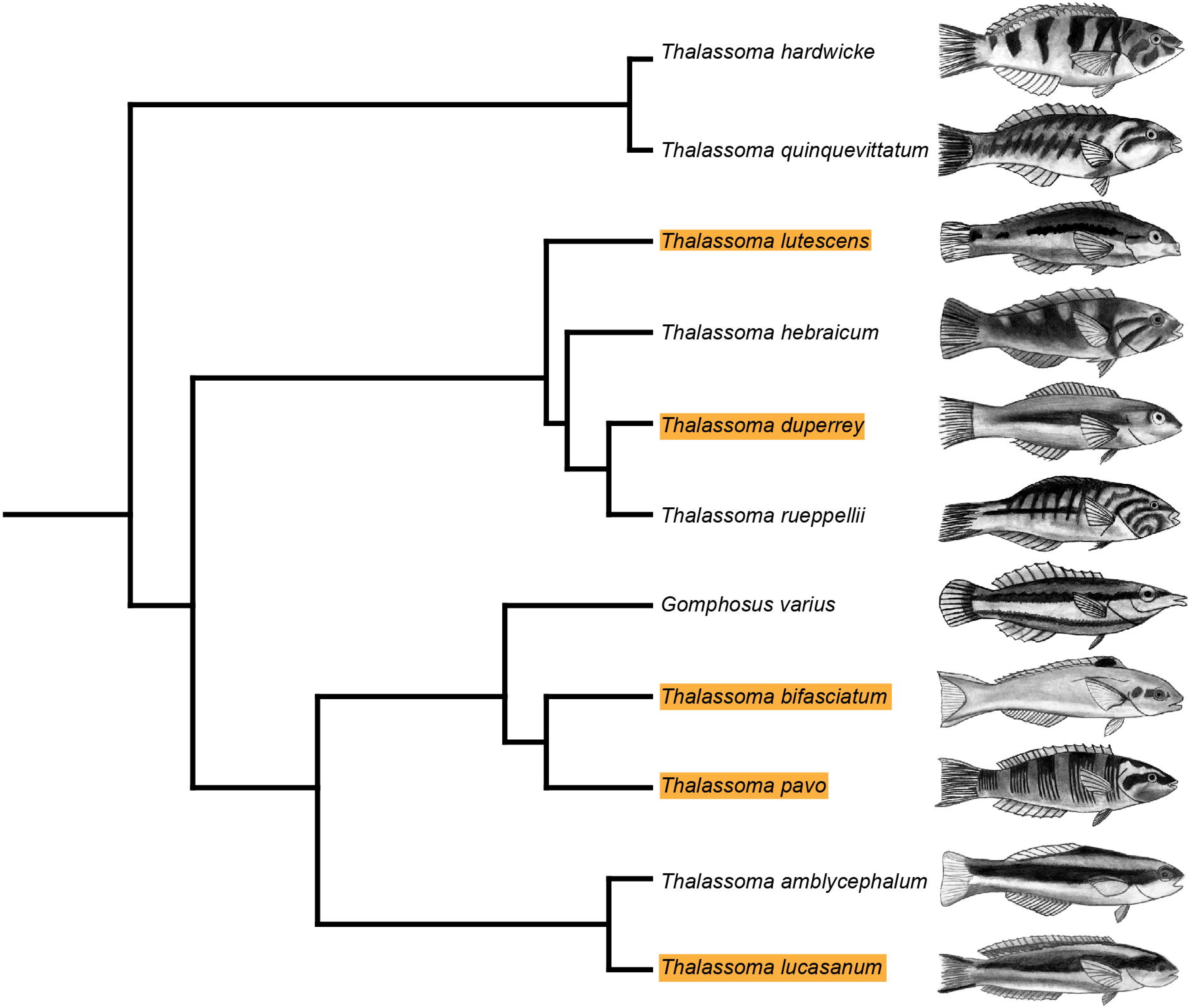
T. lucasanum	*G. varius*	1.61	2.13	--	--
	T. amblycephalum	--	1.85	1.85	--
	T. hardwicke	1.76	1.88	--	--
	T. hebraicum	1.90	1.95	--	--
	T. quinquevittatum	1.88	--	--	--
	T. rueppellii	1.87	1.92	--	1.96
T. lutescens	*G. varius*	--	--	1.89	1.75
	T. amblycephalum	1.81	--	1.68	--
	T. hardwicke	1.71	2.19	1.84	1.77
	T. hebraicum	1.82	--	--	1.65
	T. quinquevittatum	1.79	--	1.86	--
	T. rueppellii	1.67	--	1.87	1.71
T. pavo	*G. varius*	1.62	--	1.85	--
	T. amblycephalum	--	1.99	1.69	--
	T. hardwicke	1.83	1.99	1.81	--
	T. hebraicum	--	2.14	--	--
	T. quinquevittatum	--	--	1.81	--
	T. rueppellii	--	2.10	1.82	--

Table 3: Multiple regression analyses reveal the traits that contribute the most to bite force

	CAR Scores[a]									
	Residual MA of A2 Muscle[b]	Residual MA of A3 Muscle[b]	$\log_{10}$ A2 Mass	$\log_{10}$ A3 Mass	$\log_{10}$ A2 Fiber Length	$\log_{10}$ A3 Fiber Length	Adjusted R^2	df (residual)	*F*-Ratio	*p*-value
Gomphosus varius	0.04	0.09	**0.26**	0.24	0.18	0.19	0.99	10	374.68	<0.00001
Thalassoma amblycephalum	0.07	0.10	**0.27**	0.25	0.15	0.16	0.99	10	373.10	<0.00001
T. bifasciatum	0.10	0.00	**0.29**	0.26	0.15	0.18	0.99	12	270.29	<0.00001
T. duperrey	0.15	0.06	**0.26**	0.19	0.20	0.14	0.98	9	96.87	<0.00001
T. hardwicke	0.00	0.14	0.19	0.22	0.17	**0.28**	0.93	17	48.14	<0.00001
T. hebraicum	0.09	0.01	0.25	**0.26**	0.22	0.17	0.99	13	2008.12	<0.00001
T. lucasanum	0.06	0.11	0.16	**0.28**	0.13	0.25	0.97	11	72.56	<0.00001
T. lutescens	0.10	0.07	**0.22**	0.20	0.19	0.21	0.99	11	175.30	<0.00001
T. pavo	0.00	0.01	0.23	0.24	0.25	**0.28**	0.99	13	531.04	<0.00001
T. quinquevittatum	0.04	0.05	0.23	**0.33**	0.19	0.16	0.99	8	335.26	<0.00001
T. rueppellii	0.19	0.06	**0.27**	0.26	0.13	0.09	0.99	11	10664.11	<0.00001


A
1.25
1.00
0.75
0.50
0.25
$\log_{10}$ Vertical Gape Distance
1.4
1.6
1.8
2.0
$\log_{10}$ Body Length
B
1.0
0.5
0.0
$\log_{10}$ Premaxillary Protrusion Distance
1.4
1.6
1.8
2.0
$\log_{10}$ Body Length
C
0.5
0.0
−0.5
−1.0
−1.5
$\log_{10}$ Total Bite Force
1.4
1.6
1.8
2.0
$\log_{10}$ Body Length
D
1.4
1.2
1.0
0.8
Maxillary KT
1.4
1.6
1.8
2.0
$\log_{10}$ Body Length


A
Residual Mechanical Advantage of A2
0.10
0.05
0.00
−0.05
−0.10
1.4
1.6
1.8
2.0
$\log_{10}$ Body Length
B
Residual Mechanical Advantage of A3
0.10
0.05
0.00
−0.05
−0.10
1.4
1.6
1.8
2.0
$\log_{10}$ Body Length
C
$\log_{10}$ A2 Mass
−1
−2
−3
1.4
1.6
1.8
2.0
$\log_{10}$ Body Length
D
$\log_{10}$ A3 Mass
−1
−2
−3
1.4
1.6
1.8
2.0
$\log_{10}$ Body Length
E
$\log_{10}$ A2 Fiber Length
1.25
1.00
0.75
1.4
1.6
1.8
2.0
$\log_{10}$ Body Length
F
$\log_{10}$ A2 Fiber Length
1.4
1.2
1.0
0.8
1.4
1.6
1.8
2.0
$\log_{10}$ Body Length


Thalassoma hardwicke
Thalassoma quinquevittatum
Thalassoma lutescens
Thalassoma hebraicum
Thalassoma duperrey
Thalassoma rueppellii
Gomphosus varius
Thalassoma bifasciatum
Thalassoma pavo
Thalassoma amblycephalum
Thalassoma lucasanum


A
pmx
mx
A1
A3
dt
art
A2
pop
B
6
9
11
5
4
10
7
1
12
2
8
3
C
c
o
f
i
D
c
o
f
i